AD A125701

JACKKNIFING KERNEL TYPE DENSITY ESTIMATORS

J. S. Rustagi
and
S. Dynin

DEPARTMENT OF STATISTICS

The Ohio State University

OSU

MAR 16 1983

A

COLUMBUS, OHIO

A

JACKKNIFING KERNEL TYPE DENSITY ESTIMATORS

J. S. Rustagi
and
S. Dynin

Technical Report No. 280
Department of Statistics
The Ohio State University
Columbus, Ohio 43210

February, 1983

*Supported in part by Contract No. N00014-82-K-0291 (NR 042-403) by Office of Naval Research. Reproduction in whole or in part is permitted for any purpose of the United States Government.

1. Introduction

Jackknifing techniques are increasingly being applied to data analysis for bias reduction. In robust estimation, several studies have recently been published giving asymptotic properties of jackknifed estimates. Cheng (1982) has demonstrated the validity of jackknifing L-estimates under various conditions on the score function. Efron (1982) has shown that jackknife turns out to be a special case of his bootstrap technique.

In problems of density estimation, improvement of kernel type estimates was proposed by Schucany and Sommer (1977) through the technique of combining several estimates using different kernels. Usually it is possible to reduce bias in kernel-type estimates simply by a judicious choice of a kernel. However, in that case, the estimates of density functions can be negative. The situation has been described by Stute (1982) in his paper showing that the use of non-negative kernels does not allow the possibility of reduction of bias.

In this paper, the effect of jackknifing using leave-out rules, is studied. Pseudovalues in case of density estimates are defined and optimal properties of the jackknifed estimates are given. It is shown that the asymptotic behavior of the jackknifed estimates is the same as that of the classical estimate. A Berry-Esseen type central limit theorem showing the normality of the jackknifed estimate is also given.

2. Pseudovalues

Let $X_1, X_2, \ldots, X_n$ be a random sample from a population with cumulative distribution function $F(x)$ and probability density function $f(x)$. Let $K(x)$ be a given kernel function with the following properties.

P(i) $\quad \sup |K(x)| < \infty$

P(ii) $\quad \int K(x)dx = 1$

P(iii) $\quad \lim_{x\to\infty} |xK(x)| = 0$

P(iv) $\quad \int_{-\infty}^{\infty} x^i K(x)dx = 0, \; i = 1, 2, \ldots, r-1$

$\quad$ and $\int |x^r K(x)| dx < \infty$

Let $F_n(x)$ be the empirical distribution function based on the random sample and let h_n be a sequence of constants. Then the kernel density estimates of $f(x)$, defined by Rosenblatt (1956) and Parzen (1962) are given by

$$\hat{f}_n(x) = (n\, h_n)^{-1} \sum_{i=1}^{n} K\left(\frac{x-X_i}{h_n}\right)$$

$$= h_n^{-1} \int_{-\infty}^{\infty} K\left(\frac{x-y}{h_n}\right) dF_n(y) \tag{2.1}$$

It is well known that the expectation of $\hat{f}_n(x)$,

$$E[\hat{f}_n(x)] = h_n^{-1} \int K(\frac{x-y}{h_n}) dF(y) \to f(x)$$

$$\text{as } n \to \infty \text{ and } h_n \to 0,$$

Also the variance of $\hat{f}_n(x)$,

$V[\hat{f}_n(x)] \to 0$ if in addition $n\, h_n \to \infty$. The above results can be found in an extensive survey of probability density estimation by Tapia and Thompson (1978).

Let $F_{n-1}^i(x)$ is the empirical distribution function of the random sample $X_1,\ldots,X_n$ with the observation X_i removed and let

$\hat{f}_{n-1}^i(x) = \frac{1}{h_{n-1}} \int K(\frac{x-y}{h_{n-1}}) dF_{n-1}^i(y)$ where h_{n-1} is a sequence of constants based on n-1 observations.

Define the pseudovalues as follows.

$$\hat{f}_S^i(x) = \frac{h_n^{-r}}{h_n^{-r}-h_{n-1}^{-r}} \hat{f}_n(x) - \frac{h_{n-1}^{-r}}{h_n^{-r}-h_{n-1}^{-r}} \hat{f}_{n-1}^i(x) \tag{2.2}$$

The jackknifed estimate of the probability density function is then defined by the following

$$\hat{f}_J(x) = \frac{1}{n} \sum_{i=1}^{n} \hat{f}_S^i(x) \tag{2.3}$$

3. Properties of Jackknifed Estimates

In this section, several properties of the Jackknifed estimates are discussed especially its bias reduction property. The difficulty of bias reduction without negative kernels, has been demonstrated by several authors, e.g. Stute (1982, p. 419). Notice that for sufficiently smooth F's, it is always possible to reduce the bias $E[f_n(t)]-f(t)$, by choosing appropriate kernels K. Also among the class of nonnegative K's, P(iv) can be achieved only for $r = 1$, thus giving h_n^2 as the best possible error rate. For better results, one has to include also these K's for which $K(y)$ may be negative, leading to an estimate of $\hat{f}_n(t)$ which may be negative.

From now on, we shall assume that kernels satisfy the following additional properties,

P(v) The rth derivatives of density function satisfy a Lipshitz condition

$$|f^{(r)}(x)-f^{(r)}(y)| < c\ |x-y|^{\alpha}, \quad 0 \leq \alpha \leq 1$$

for all x, and y.

P(vi) $\{h_n\}$ is a sequence of constants satisfying

$$\frac{h_n}{h_{n-1}} = 1 + o(1),$$

P(vii) $\int |x^{r+\alpha}\ K(x)|\,dx < \infty.$

We prove the following theorem.

Theorem 3.1. Under the conditions P(i) - P(vii) on kernel K(x),

$$E[\hat{f}_n(x)] = f(x) + \frac{h_n^r f^{(r)}(x) \int_{-\infty}^{\infty} z^r K(-z)dz}{r!} + O(h_n^{r+\alpha}) .$$

Proof.

$$E[\hat{f}_n(x)] = \int_{-\infty}^{\infty} K(-z)f(x+zh_n)dz$$

$$= f(x) + \int_{-\infty}^{\infty} K(-z)[f(x+zh_n)-f(x)]dz. \tag{3.1}$$

Using the Taylor's expansion with integral remainder, we have,

$$f(x+zh_n) = f(x) + zh_n f'(x) + \ldots + \frac{z^r h_n^{r-1}}{(r-1)!} f^{(r-1)}(x) + \int_x^{x+zh_n} \frac{(x+zh_n-\xi)^{r-1}}{(r-1)!} f^{(r)}(\xi)d\xi. \tag{3.2}$$

Using P(iv), we have

$$E[\hat{f}_n(x)] = f(x) + \int_{-\infty}^{\infty} \int_x^{x+zh_n} \frac{(x+zh_n-\xi)^{t-1}}{(r-1)!} f^{(r)}(\xi)K(-z)d\xi dz \tag{3.3}$$

Let

$$\int_{-\infty}^{\infty} \int_x^{x+zh_n} \frac{(x+zh_n-\xi)^{r-1}}{(r-1)!} [f^{(r)}(\xi)-f^{(r)}(x)+f^{(r)}(x)]K(-z)d\xi dz$$

$$= \frac{f^{(r)}(x)}{r!} \int_{-\infty}^{\infty} K(-z)(zh_n)^r d_z +$$

$$\int_{-\infty}^{\infty} \int_x^{x+zh_n} \frac{(x+zh_n-\xi)^{r-1}}{(r-1)!} [f^{(r)}(\xi)-f^{(r)}(x)]K(-z)d\xi dz \tag{3.4}$$

The second integral on the right of (3.4) in absolute value is

$$\leq c \int_{\infty}^{\infty} \int_{x}^{x+zh_n} \frac{(x+zh_n-\xi)^{r-1}}{(r-1)!} |\xi-x|^{\alpha} K(-z)d\xi dz \text{ using } P(v)$$

$$\leq c \int_{\infty}^{\infty} |K(-z)||zh_n|^{\alpha} \int_{x}^{x+zh_n} \frac{(x+zh_n-\xi)^{r-1}}{(r-1)!} d\xi\, dz$$

$$\leq \frac{c}{r!} \int_{\infty}^{\infty} |K(-z)||zh_n|^{r+\alpha} dz \quad \text{by } P(vii)$$

$$= O(h^{r+\alpha}). \tag{3.5}$$

Combining (3.4) and (3.5), we get the result (3.1) proving the theorem. $\square$

Using the results of Theorem 3.1, we can find the bias of the jackknifed estimate $\hat{f}_J(x)$. The result is given in Theorem 3.2.

Theorem 3.2. Under the conditions P(i) - P(vii) for the kernel, the bias of the jackknifed estimate $\hat{f}_J(x)$ is

$$O\left(\frac{h_n^{r+1}}{(h_{n-1}-h_n)^{1-\alpha}}\right) \tag{3.6}$$

Proof. The expectation of the jackknifed estimate, from equation (2.3), is given by

$$E[\hat{f}_J(x)] = (h_n^{-r}-h_{n-1}^{-r})^{-1}\{h_n^{-r}E[\hat{f}_n(x)]$$

$$- h_{n-1}^{-r}E[\hat{f}_{n-1}^{i}(x)]\}$$

$$\text{Bias } \hat{f}_J(x) = E(\hat{f}_J(x)) - f(x)$$

$$= (h_n^{-r} - h_{n-1}^{-r})^{-1}\{h_n^{-r} \text{ Bias } \hat{f}_n(x)$$

$$-h_{n-1}^{-r} \text{ Bias } \hat{f}_{n-1}^{i}(x)\} \tag{3.7}$$

Using (3.3), we have

$$\text{Bias } \hat{f}_J(x) = (h_n^{-r}-h_{n-1}^{-r})^{-1}\{h_n^{-r} \int_{-\infty}^{\infty} \int_{x}^{x+zh_n} \frac{(x+zh_n-\xi)^{r-1}}{(r-1)!} f^{(r)}(\xi)K(-z)d\xi\, dz$$

$$- h_{n-1}^{-r} \int_{-\infty}^{\infty} \int_{x}^{x+zh_{n-1}} \frac{(x+zh_{n-1}-\xi)^{r-1}}{(r-1)!} f^{(r)}(\xi)K(-z)d\xi dz \tag{3.8}$$

Making the transformation

$$\xi - x = z\, h_n\, \eta$$

in the first integral and $\xi - x = zh_{n-1}\eta$ in the second integral in (3.8), we have

$$\text{Bias } \hat{f}_J(x) = \frac{(h_n^{-r}-h_{n-1}^{-r})^{-1}}{(r-1)!}$$

$$\int_{\infty}^{\infty}\int_0^1 (1-\eta)^{r-1}[f^{(r)}(x+\eta\, h_n z)-f^{(r)}(x+\eta h_{n-1}z)]K(-z)z^r d\eta\, dz$$

$$\leq \frac{(h_n^{-r}-h_{n-1}^{-r})^{-1}}{(r-1)!} \int_{\infty}^{\infty}\int_0^1 (1-\eta)^{r-1}\eta^{\alpha}(h_n-h_{n-1})^{\alpha} z^{\alpha+r}K(-z)d\eta\, dz$$

$$= O\left(\frac{(h_n-h_{n-1})^{\alpha}}{h_n^{-r}-h_{n-1}^{-r}}\right)$$

Assuming $\frac{h_{n-1}-h_n}{h_n} = o(1)$, we have the bias reduced to

$$O\left(\frac{h_n^{r+1}}{(h_{n-1}-h_n)^{1-\alpha}}\right)$$

since $$\left(\frac{h_n}{h_{n-1}}\right)^r = \left(1 - \frac{h_{n-1}-h_n}{h_n}\right)^r$$

$$\doteq 1 - r\left(\frac{h_{n-1}-h_n}{h_n}\right)$$

Remark. The comparison of theorem 3.2 with theorem 3.1 clearly demonstrates the reduction of bias of the jackknifed estimate $\hat{f}_J(x)$ by at least of the term h_n^r. By the proper choice of h_n, we can reduce the second term also under certain smoothness conditions on the probability density function f.

Variance of $\hat{f}_J(x)$

Using the expressions for jackkifed estimate in (2.2) and (2.3) we have the variance of the estimate, σ_J^2 as follows:

$$\sigma_J^2 = n^{-1}(h_n^{-r}-h_{n-1}^{-r})^{-2}\ \mathrm{Var}\{h_n^{-r-1}K(\frac{x-y}{h_n})$$

$$- h_{n-1}^{-r-1}K(\frac{x-y}{h_{n-1}})\}$$

$$= A + B$$

where

$$A = n^{-1}(h_n^{-r}-h_{n-1}^{-r})^{-2}\{\int h_n^{-r-1}[K(\frac{x-y}{h_n})$$

$$- h_{n-1}^{-r-1}K(\frac{x-y}{h_{n-1}})]^2 f(y)dy$$

$$B = n^{-1}(h_n^{-r}-h_{n-1}^{-r})^{-2}[\int h_n^{-r-1}[k(\frac{x-y}{h_n})$$

$$- h_{n-1}^{-r-1}K(\frac{x-y}{h_{n-1}})f(y)dy]^2$$

Notice that with $z = (x-y)h_n^{-1}$, we have

$$A = n^{-1}(h_n^{-r}-h_{n-1}^{-r})^{-2}h_n^{-2r-1}$$

$$\int_{-\infty}^{\infty}[K(z)-(\frac{h_n}{h_{n-1}})^{r+1}K(z\ \frac{h_n}{h_{n-1}})]^2 f(x-zh_n)dz$$

$$= (n\ h_n)^{-1}(h_n^{-r}-h_{n-1}^{-r})^{-2}h_n^{-2\gamma}(1-\frac{h_n}{h_{n-1}})^2$$

$$\int_{\infty}^{\infty}\{[\frac{K(z)-K(z\ \frac{h_n}{h_{n-1}})}{1-\frac{h_n}{h_{n-1}}}] - \frac{K(z\ \frac{h_n}{h_{n-1}})}{1-\frac{h_n}{h_{n-1}}}\ [(\frac{h_n}{h_{n-1}})^{r+1}-1]\}^2$$

$$f(x - z\ h_n)dz.$$

In the limit when $\frac{h_n}{h_{n-1}} \to 1$, $n\, h_n \to \infty$, $h_n \to 0$ as $n \to \infty$,

$$\lim_{n\to\infty}(n\, h_n A) = f(x) \int_{-\infty}^{\infty} (z\, K'(z)+K(z)(r+1))^2 dz\,,$$

so that $$A = \frac{f(x)}{n\, h_n} \int_{-\infty}^{\infty} [z\, K'(z) + K(z)(r+1)]^2 dz + 0(\frac{1}{n\, h_n})\,.$$

We use the conditions that $\int z^2[K'(z)]^2 dz < \infty$ and $\int_m^{\infty} K^2(z) = 0(m^{-2})$
Also note that from theorem (3,2), we have

$$B = \frac{1}{n}[f(x) + 0(\frac{h_n^{r+1}}{(h_{n-1}-h_n)^{1-\alpha}})^2$$

Hence, we have, as $n \to \infty$,

$$\sigma_J^2 = \frac{f(x)}{n\, h_n} \int_{-\infty}^{\infty} \{z\, K'(z) + (r+1)K(z)\}^2 dz$$

$$+ o(\frac{1}{n\, h_n})$$

as B contains terms of much lower order than $o(n^{-1}h_n^{-1})$.

Notice that $\sigma_J^2 > 0$ since $zK'(z) + (r+1)K(z) \nmid 0$ for all integrable functions $K'(z)$. If $zK'(z) + (r+1)K(z) \equiv 0$, then $K(z) = z^{-(r+1)}$ which is not integrable.

4. Central Limit Theorem for the Jackknifed Estimate

Since the estimate is a sum of n independently distributed random variables, the following theorem can be proved under the usual conditions of central limit theorems since $\hat{f}_n(x)$ and similarly $\hat{f}^i_{n-1}(x)$ are asymptotically normal.

Theorem. As $n \to \infty$,

$$Pr\left\{\frac{\hat{f}_J(x)-f(x)}{\sigma_J} \le y\right\} \to \Phi(y)$$

To find the Berry-Esseen bounds, we need $(2+\delta)$-th moment of the jackknifed estimate (2.2) which is an average of $\hat{f}^i_S(x)$, $i=1,2,\ldots,n$ given by (2.1). Using the expresssion s in terms of kernel K for $\hat{f}^{(i)}_n(x)$ and $\hat{f}^i_{n-1}(x)$, we can write the $(2+\delta)$-th moment of $\hat{f}_J(x)$. By Jensen's inequality, we have

$$\mu^i_{2+\delta} = E|\hat{f}^i_S(x)-E\hat{f}^i_S(x)|^{2+\delta} \le A_{2+\delta}\{E|\hat{f}^i_S(x)|^{2+\delta}$$

$$+ |E(f^i_S(x))|^{2+\delta}\}$$

$$\le 2\ A_{2+\delta}E|f^i_S(x)|^{2+\delta}\}$$

where $A_{2+\delta}$ is constant depending on δ and

$$(E|\hat{f}^i_S(x)|)^{2+\delta} \le E|\hat{f}^i_S(x)|^{2+\delta}.$$

Hence $\mu_{2+\delta} = \Sigma\mu^i_{2+\delta} \le 2nA_{2+\delta}\ E|\hat{f}^i_S(x)|^{2+\delta}$

Now

$$E|f_S^i(x)|^{2+\delta} = n^{-(2+\delta)}(h_n^{-r}-h_{n-1}^{-r})^{-(2+\delta)}.$$

$$\cdot \int [h_n^{-r-1} K(\frac{x-y}{h_n}) - h_{n-1}^{-r-1}K(\frac{x-y}{h_{n-1}})]^{2+\delta} f(y)dy$$

$$= n^{-(2+\delta)}(h_n^{-r}-h_{n-1}^{-r})^{-(2+\delta)} h_n^{-(r+1)(2+\delta)+1} .$$

$$\cdot \int [K(z) - (\frac{h_{n-1}}{h_n})^{-(r+1)}K(z \frac{h_n}{h_{n-1}})]^{2+\delta} f(x-z\ h_n)dz$$

$$\leq c\ n^{-(2+\delta)}(h_n^{-r}-h_{n-1}^{-r})^{-(2+\delta)}h_n^{-(r+1)(2+\delta)+1} .$$

Since $\int [K(z) - (\frac{h_{n-1}}{h_n})^{-(r+1)}K(z\frac{h_n}{h_{n-1}})]^{2+\delta} f(x-zh_n)dz$

$$\leq C(1 - \frac{h_n}{h_{n-1}})^{2+\delta}$$

Hence

$$\mu_{2+\delta} \leq 2n\ A_{2+\delta}\ C.\ n^{-(2+\delta)}h_n^{-1-\delta}(\frac{h_n}{h_{n-1}} - 1)^{-(2+\delta)}$$

$$\leq \frac{C}{n^{1+\delta}\ h_n^{1+\delta}} ,$$

Giving

$$\mu_{2+\delta} = O(\frac{1}{(n\ h_n)^{1+\delta}}).$$

For estimation of $\mu_{2+\delta}$, we use the conditions,

$$\int_{-\infty}^{\infty} z^{2+\delta}(K'(z))^{2+\delta} dz < \infty$$

$$\int_{m}^{\infty} K^{2+\delta}(z)dr = o(\frac{1}{m^{2+\delta}})$$

Now we state the Berry-Esseen type theorem for $f_J(x)$. For reference see Chao and Teicher (1978, p. 299).

Theorem 4.1

$$\sup|P\{\frac{n(\hat{f}_J(x)-E\hat{f}_J(x))}{\sigma_J} < x\} - \Phi(x)| \leq$$

$$\leq C_\delta \frac{\mu_{2+\delta}}{\sigma_J^{2+\delta}}$$

The above result gives the uniform convergence of the central limit theorem for the jackknifed density estimate.

References

[1] Cheng, Kuang-Fu (1982). Jackknifing L-Estimates, Canadian J. Statist. 10, 49-58.

[2] Chow, Y. S. and Teicher, H. (1978). Probability Theory. Springer-Verlag, New York.

[3] Efron, Bradley (1982). The Jackknife, the Bootstrap and Other Resampling Plans, Society of Industrial and Applied Mathematics, Philadelphia, Pennsylvania.

[4] Parzen, E. (1962). On estimation of a probability-density function and its mode, Ann. Math. Statist. 33, 1065-76.

[5] Rosenblatt, M. (1956). Remarks on some nonparametric estimates of a density function, Ann. Math. Statist. 26, 832-837.

[6] Schucany, W. R. and Sommer, John P. (1977). Improvement of kernel type density estimators, J. Amer. Statist. Assoc. 72, 420-423.

[7] Stute, Winfred (1982). A law of the logarithm for kernel density, Ann. Prob. 10, 414-422.

[8] Tapia, R. A. and Thompson, J. R. (1978). Nonparametric Probability Density Estimation, Johns Hopkins University Press, Baltimore, Maryland.

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO. AD-A125 701	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) JACKKNIFING KERNEL TYPE DENSITY ESTIMATORS		5. TYPE OF REPORT & PERIOD COVERED Technical Report - #280 4/1/80-2/28/83
		6. PERFORMING ORG. REPORT NUMBER 763099/714531
7. AUTHOR(s) J. S. Rustagi and S. Dynin		8. CONTRACT OR GRANT NUMBER(s) N00014-82-K-0291
9. PERFORMING ORGANIZATION NAME AND ADDRESS The Ohio State University Research Foundation, 1314 Kinnear Road Columbus, Ohio 43212		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR 042-403
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Department of the Navy Arlington, Virginia 22207		12. REPORT DATE February, 1983
		13. NUMBER OF PAGES 14
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Jackknifing method, density estimates, kernel estimates

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Effect of jackknifing using leave out rules on kernel density estimates is studied. Pseudovaiues are defined and the optimal properties of jackknifed estimates are given. It is shown that the asymptotic behavior of the jackknifed estimates is the same as that of the classical estimate. A Berry--Esseen type central limit theorem is also given for these estimates.